

SAM/DAG/Job Submission Integration

Dennis Box, FNAL REX/CD

8/14/12

Using Jobsub with SAM

- Jobsub_tools : A KITS product with jobsub, (experiment)_jobsub, dagNabbit, condor_wrappers, etc.
- IFDHC: A KITS product with nice shell and python interfaces to SAMWeb
- General Idea:
 - setup jobsub_tools
 - jobsub (most of same arguments you would use with minerva_jobsub)
 - setup ifdhc
 - ifdh startProject (SAMProjectName)
 - While (more_files_in_project) { Ifdh getNextFile; process_file() }
 - Ifdh endProject

Iteration 0:

- Use DAG generator `dagNabbit.py` to shepherd previously described steps through a simple workflow
- PROS: it works
- CONS: its messy, see
 - https://cdcv.sfnal.gov/redmine/projects/ifront/wiki/Integrating_DAGS_with_SAM_
 - users asked for more convenient interface

Iteration 1:

- (Brian Tice:) Why not put options into jobsub that generate DAG, start SAM, etc.
- Jobsub v0_6 implements this
 - Jobsub -h lists options by functional groups, the SAM Options group:
SAM Options:
`--dataset_definition=DATASET_DEFINITION`
SAM dataset definition used in a Directed Acyclic
Graph (DAG)
`--project_name=PROJECT_NAME`
optional project name for SAM DAG

Usage Example:

a sample script

- Environment variables generated and populated by jobsub script:
 - EXPERIMENT
 - SAM_PROJECT_NAME
- Ifdh findProject, getNextFile, fetchInput, updateFileStatus used in this example

```
$ cat consume_sam.sh
#!/bin/sh -x
source /grid/fermiapp/products/$EXPERIMENT/etc/setups.sh
setup ifdhc

CPURL=`ifdh findProject $SAM_PROJECT_NAME $EXPERIMENT`
HOST=`/bin/hostname`
WHO=`/usr/bin/whoami`
CONSUMER_ID=`ifdh establishProcess $CPURL demo 1 $HOST $WHO`
FURI=`ifdh getNextFile $CPURL $CONSUMER_ID`
while [ "$FURI" != "" ]
do
  FNAME=`ifdh fetchInput $FURI `
  ifdh updateFileStatus $CPURL $CONSUMER_ID $FNAME transferred
  sleep 1
  ifdh updateFileStatus $CPURL $CONSUMER_ID $FNAME consumed
  rm -f $FNAME
  FURI=`ifdh getNextFile $CPURL $CONSUMER_ID`
done
```

Usage Example:

submitting the script in a DAG

```
$ source /grid/fermiapp/products/minerva/etc/setups.sh
$ setup jobsub_tools

$ jobsub --dataset_definition mwm_test_2 -N 4 ./consume_sam.sh
/minerva/app/users/condor-tmp/dbox/consume_sam.sh_20120814_110640_21721_1.dag
submitting....
Submitting job(s).
Logging submit event(s).
1 job(s) submitted to cluster 3997196.
```

```
File for submitting this DAG to Condor      : /minerva/app/users/condor-
tmp/dbox/consume_sam.sh_20120814_110640_21721_1.dag.condor.sub
Log of DAGMan debugging messages      : /minerva/app/users/condor-
tmp/dbox/consume_sam.sh_20120814_110640_21721_1.dag.dagman.out
Log of Condor library output      : /minerva/app/users/condor-
tmp/dbox/consume_sam.sh_20120814_110640_21721_1.dag.lib.out
Log of Condor library error messages      : /minerva/app/users/condor-
tmp/dbox/consume_sam.sh_20120814_110640_21721_1.dag.lib.err
Log of the life of condor_dagman itself      : /minerva/app/users/condor-
tmp/dbox/consume_sam.sh_20120814_110640_21721_1.dag.dagman.log
```

Usage Example:

submitting the script in a DAG

```
$ source /grid/fermiapp/products/minerva/etc/setups.sh
$ setup jobsub_tools

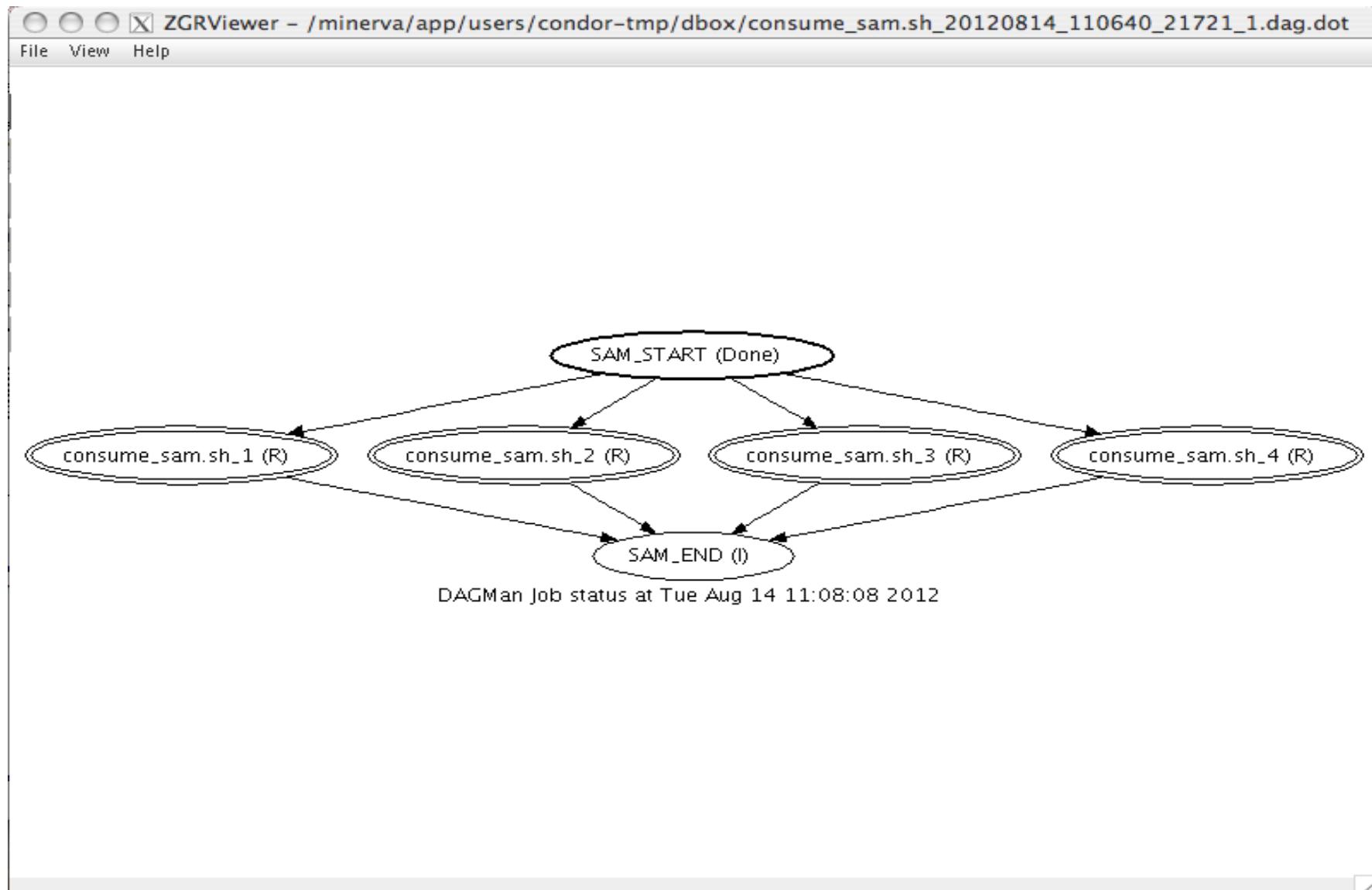
$ jobsub --dataset_definition mwm_test_2 -N 4 ./consume_sam.sh
/minerva/app/users/condor-tmp/dbox/consume_sam.sh_20120814_110640_21721_1.dag
submitting....
Submitting job(s).
Logging submit event(s).
1 job(s) submitted to cluster 3997196.
```

```
File for submitting this DAG to Condor      : /minerva/app/users/condor-
tmp/dbox/consume_sam.sh_20120814_110640_21721_1.dag.condor.sub
Log of DAGMan debugging messages      : /minerva/app/users/condor-
tmp/dbox/consume_sam.sh_20120814_110640_21721_1.dag.dagman.out
Log of Condor library output      : /minerva/app/users/condor-
tmp/dbox/consume_sam.sh_20120814_110640_21721_1.dag.lib.out
Log of Condor library error messages      : /minerva/app/users/condor-
tmp/dbox/consume_sam.sh_20120814_110640_21721_1.dag.lib.err
Log of the life of condor_dagman itself      : /minerva/app/users/condor-
tmp/dbox/consume_sam.sh_20120814_110640_21721_1.dag.dagman.log
```

Usage Example:

Monitoring Progress:

```
$ /grid/fermiapp/common/graphviz/zgrviewer/zgrview /minerva/app/users/condor-tmp/dbox/consume_sam.sh_20120814_110640_21721_1.dag.dot
```



Usage Example:

Monitor Progress

Also monitor progress at http://d0dbweb.fnal.gov:8080/station_monitor/minerva/stations/



Minerva SAM station monitoring

Station	Projects	Last updated	Last activity
minerva	1	2012-08-14 11:37:14	project started at 2012-08-14 11:07:57 (29 minutes ago)

Whats Next?

- Better error handling
- User testing and feedback